Automatic Geo-referencing of Flickr Videos

Pascal Kelm
Kelm@nue.tu-berlin.de
Communication Systems Group
Technische Universität Berlin

Thursday, 06 October 2011
Task Description

Placing Task @ MediaEval Benchmark
Where in the world is it?

Castello del Valentino
Video Distribution

Flickr Database:
5125 trainings videos
5091 test videos
Image Distribution

Flickr Database: 3.6 million training images
The city of Berlin is a beautiful place.

Berlin → Germany

GEO Lookup using Gazetteers

Geographical Boundaries

GEO Location

Meta Data → Natural Language Processing → GEO Lookup using Gazetteers

Key Frames → Feature Extraction & Normalization → Nearest Neighbour Classification

Kelm: Automatic Geo-referencing of Flickr Videos
這是我上次去巴黎。在那裡，我得到了我的城堡在迪士尼樂園看。...
This was my last trip to Paris. I visited the castle in Disneyland…

Which words gives us information? Tags?
- Trip, Paris, Castle, Disneyland

Which of these nouns have got geographical information?
- Paris, Disneyland
Geographical Ambiguity

\[ c_{detected} = \arg \max \left( \sum_{j=0}^{N-1} R_j(c_0) \right) \]

\[ \sum_{j=0}^{N-1} R_j(c_M) \]

\( R(c_i) = \) Rank sum
\( c_i = \) Countries
\( N = \) Number of toponym

- Paris
  - France
  - Canada
  - Puerto Rico

- Disneyland
  - China
  - USA
  - France

Kelm: Automatic Geo-referencing of Flickr Videos
The **city** of **Berlin** is a beautiful **place**.

Berlin ➔ Germany

**Framework**

Meta Data ➔ Natural Language Processing ➔ GEO Lookup using Gazetteers

Key Frames ➔ Feature Extraction & Normalization ➔ Nearest Neighbour Classification ➔ GEO Location

The city of Berlin is a beautiful place.
World Map Segmentation

- World map is segmented in $360 \times 180$ blocks

Visual Features:
1. Color and Edge Directivity Descriptor
2. Gabor
3. Fuzzy Color and Texture Histogram
4. Color Histogram
5. Scalable Color
6. Auto Color Correlogram
7. Tamura
8. Edge Histogram
9. Color Layout
– Returns the visually most similar areas, which are represented by a mean feature vector of all training images and videos of the respective area
Visual Features

Key frame

Probability map based on visual features
Accuracy/Geographical Error

Accuracy [%] vs. Margin of error [km]

- Geographical Boundary Extraction + Visual Region Model + Visual Nearest Neighbour
- Visual Region Model + Visual Nearest Neighbour
MediaEval is part of the ICT project OpenSEM and the FP7 NoE PetaMedia

www.opensem.eu

(Data Center → MediaEval)